PERBANDINGAN KUASA UJIAN KEBAGUSAN PENYUAIAN: KES TABURAN NORMAL

Ahmad Shukri Yahaya
Pusat Pengajian Kuantitatif, UKM.

Pengenalan

Salah satu daripada penggunaan ujian kebagusan penyuaian adalah berkenaan dengan pengujian terhadap taburan bagi sesuatu sampel. Jika ujian itu adalah mengenai persetujuan di antara taburan nilai-nilai daripada sampel dengan taburan populasinya, maka ujian itu digelar ujian kebagusan penyuaian.

Ujian khi-kuasadua (Pearson 1900) yang selalu digunakan untuk menguji ujian kebagusan penyuaian hanya sesuai untuk saiz sampel yang besar sahaja. Enam jenis statistik ujian yang berdasarkan kepada taburan empiriknya dan sesuai untuk sampel saiz kecil dan besar akan digunakan. Kertas ini akan membandingkan kuasa bagi enam statistik ujian berdasarkan kepada beberapa taburan alternatif yang berlainan. Kaedah Monte Carlo (Naylor 1966) digunakan untuk menjanakan variat rawak yang diperlukan untuk menguji kuasa setiap ujian.

Statistik ujian yang digunakan

Katakan $X = (x_1, x_2, x_3, \ldots\ldots\ldots\ldots, x_n)$ adalah cerapan tak bersandar bersaiz n bagi pembolehubah rawak dengan fungsi taburan $F_0(x)$ yang tidak diketahui. Andaikan hipotesis yang hendak diuji adalah seperti berikut.

$$H_0 : F_0 = F(x), \text{ -------------------------- } (2.1)$$

untuk setiap x, iaitu sampel dikutip daripada populasi normal F(x) yang parameternya tidak diketahui.

Pengujian terhadap (2.1) digelar ujian kebagusan penyesuaian. Statistik ujian yang digunakan adalah seperti berikut:

Statistik Ujian	Tanda	Rumusan
Kolmogorov-Smirnov	D_n^+	$\underset{1 \le i \le n}{\text{maks}} \{i/n - Z_i\}$
	D_n^-	$\underset{1 \le i \le n}{\text{maks}} \{Z_i - (i - 1)/n\}$
	D_n	$\underset{1 \le i \le n}{\text{maks}} \{D_n^+ , D_n^-\}$
Anderson Darling	A_n^2	$-\{\sum_{i=1}^{n} (2i - 1) (\ln Z_i + \ln(1 - Z_{n+1-i})\} /n^{-n}$
Cramer Von Mises	W_n^2	$\sum_{i=1}^{n} \{Z_i - (2i - 1)/2n\}^2 + 1/12n$
Watson	U_n^2	$W_n^2 - n(\bar{Z} - 1/2)^2$ dengan $\bar{Z} = \sum_{i=1}^{n} Zi/n$
Kuiper	V_n	$D_n^+ + D_n^-$

Stephens 1974 dan (1970) telah mencadangkan pengubahsuaian terhadap statistik di atas supaya satu ujian ringkas bagi nilai genting didapati. Statistik yang telah diubahsuaikan ini serta statistik Kolmogorov-Smirnov akan digunakan untuk kajian seterusnya.

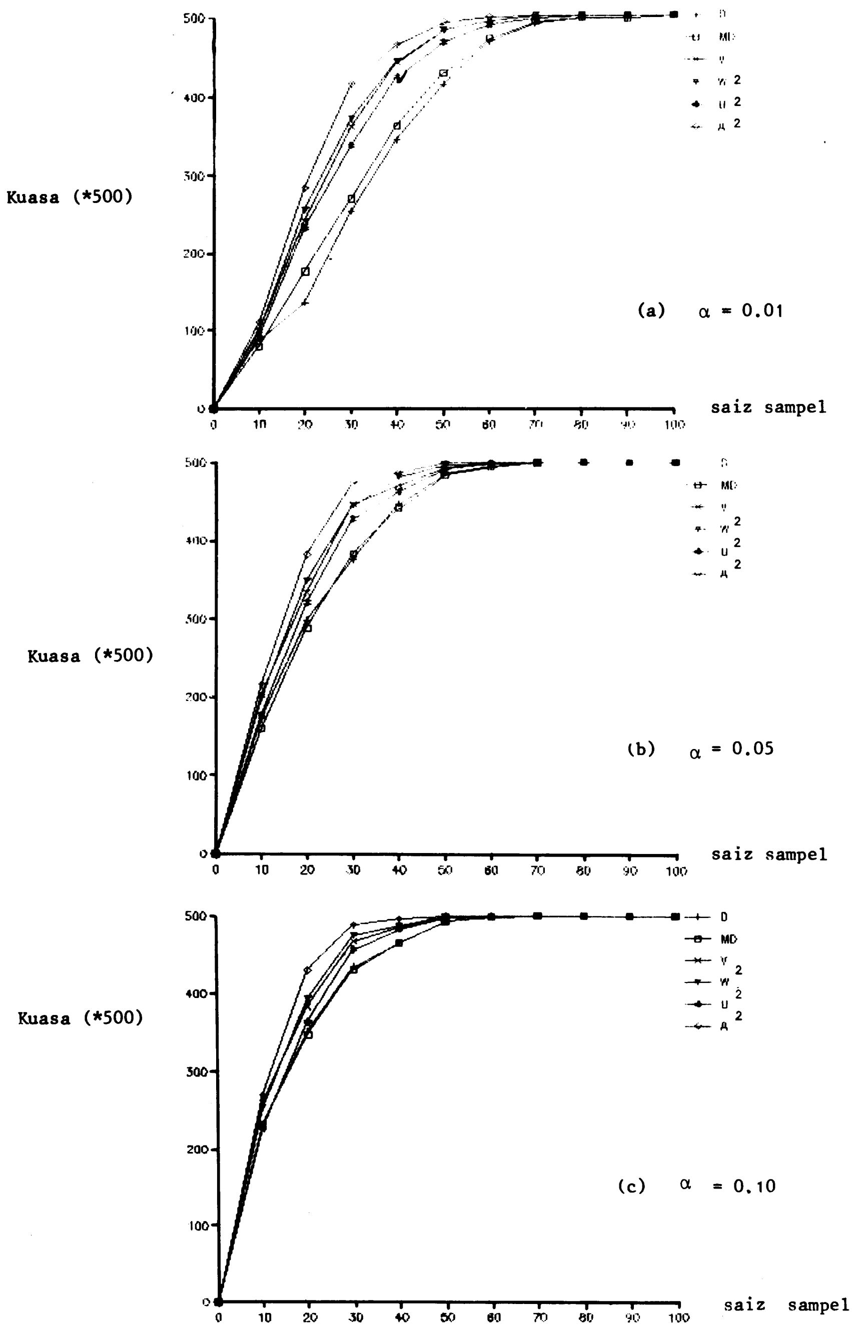
Perlaksanaan Ujian

Kuasa bagi setiap statistik ujian digunakan untuk membandingkan kecekapan statistik berkenaan. Kuasa bagi sesuatu ujian itu ditakrifkan sebagai keupayaan (disukat sebagai kebarangkalian) ujian tersebut untuk menghasilkan keputusan yang betul. Taburan alternatif yang digunakan adalah taburan seragam selanjar, taburan eksponen dan taburan khi-kuasadua dengan 5 darjah kebebasan.

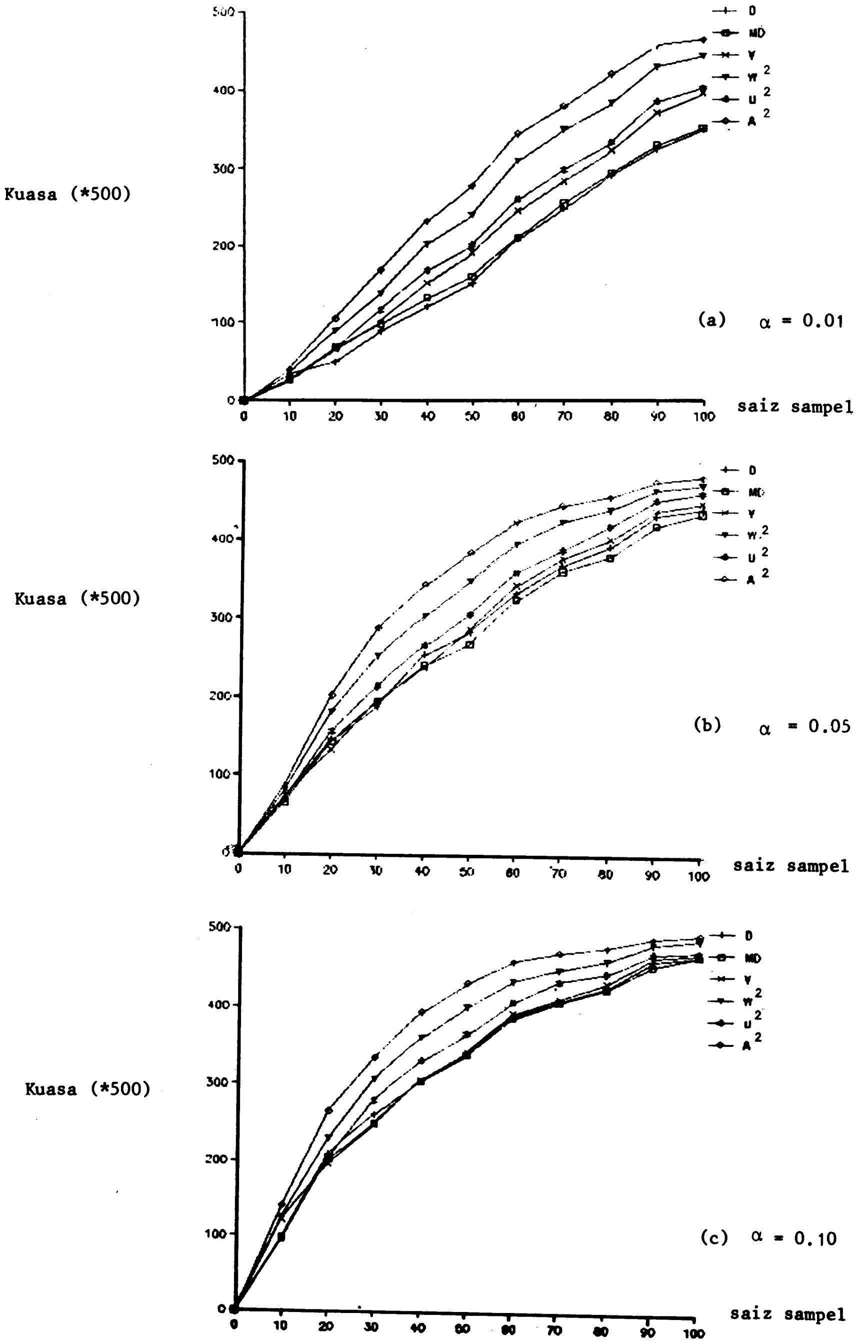
Saiz sampel yang diambil adalah dari 100. Untuk setiap saiz sampel bilangan cerapan yang digunakan adalah 500.

Keputusan dan Perbincangan

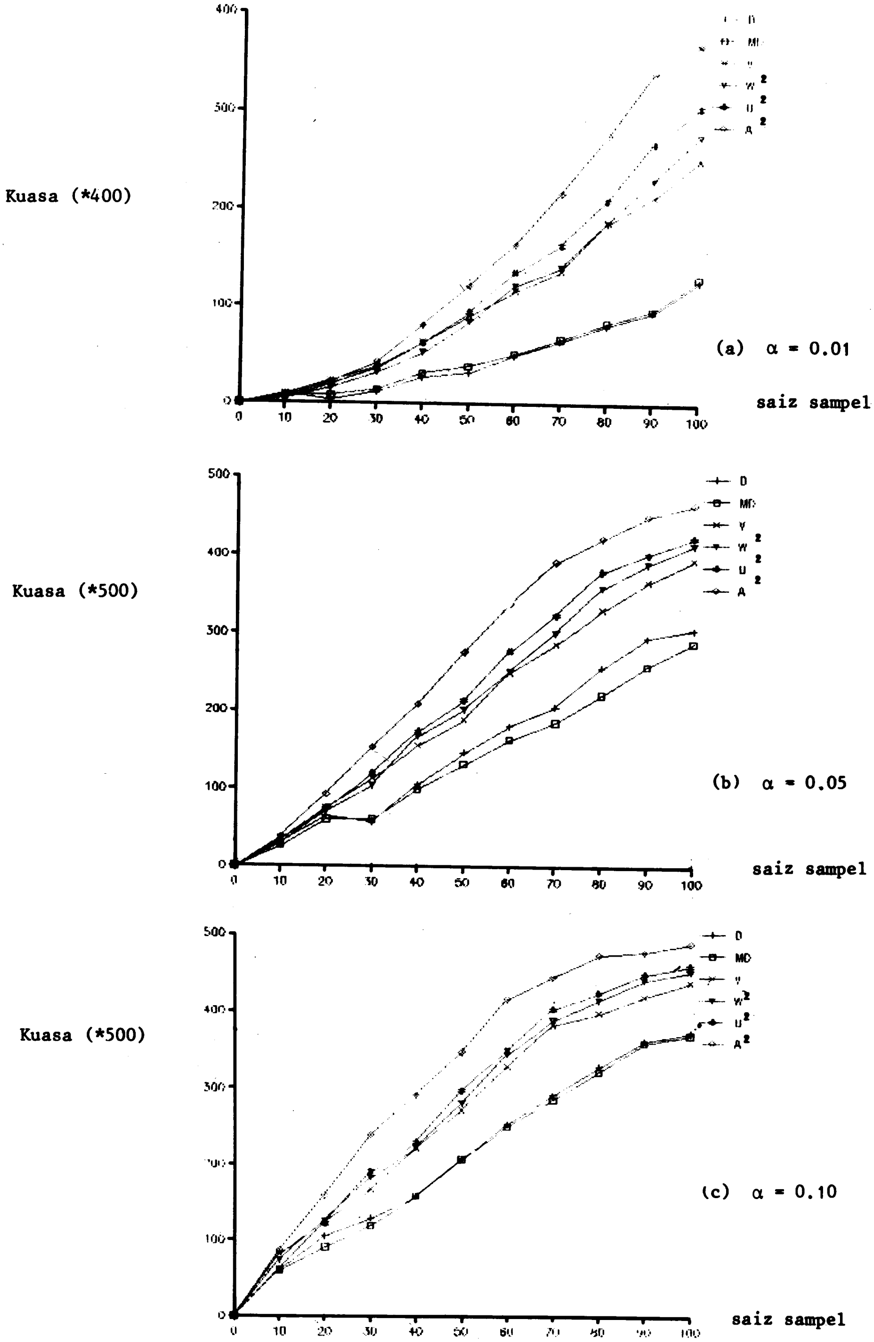
Kuasa untuk setiap statistik ujian yang berdasarkan kepada tiga taburan alternatif yang berlainan itu dilakarkan di dalam bentuk carta garis. Rajah 1 hingga 3 memberikan perbandingan kuasa untuk setiap taburan alternatif itu. Kaedah Monte Carlo digunakan untuk mendapat kuasa setiap statistik ujian itu.



RAJAH 1: KUASA UJIAN APABILA TABURAN ALTERNATIF ADALAH EKSPONEN



RAJAH 2: KUASA UJIAN APABILA TABURAN ALTERNATIF ADALAH KHI-KUASA DUA DENGAN 5 DARJAH KEBEBASAN



RAJAH 3: KUASA UJIAN APABILA TABURAN ALTERNATIF ADALAH SERAGAM

Dari Rajah 1 hingga 3, beberapa keputusan mengenai ujian yang patut digunakan untuk saiz sampel tertentu boleh diperolehi. Jadual 1 dan 2 memberikan cadangan itu apabila taburan alternatif adalah taburan seragam selanjar dan taburan eksponen. Apabila taburan alternatif adalah khi-kuasadua dengan 5 darjah kebebasan, statistik Ajne (A^2) patut digunakan. Dari jadual 1 dengan $\alpha = 0.05$ dan apabila saiz sampel adalah di antara 40 – 100 kesemua statistik ujian boleh digunakan. Walau bagaimanapun, masa komputer serta penggunaan ujian statistik yang mudah patut dipertimbangkan.

Penutup

Pada setiap taburan alternatif, statistik Ajne (A^2) memberikan kuasa terbaik. Statistik Cramer-von-Mises (W^2) menduduki tempat kedua manakala statistik Watson (U^2) berupaya menduduki tempat ketiga. Statistik Kolmogorov-Smirnov (D dan MD) tidak berupaya memberikan kuasa yang baik bagi setiap taburan alternatif.

JADUAL 1
Cadangan Statistik Ujian Yang Patut Digunakan
Apabila Taburan Alternatif Adalah Seragam (Ditandakan Dengan*)

Paras Keertian	$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$
Saiz Sampel Statistik Ujian	10–30	40–100	10	20–100	10–100
D			*		
MD			*		
V	*		*		
W^2	*		*		
U^2			*		
A^2	*	*	*	*	*

JADUAL 2

Cadangan Statistik Ujian Yang Patut Digunakan Apabila Taburan Alternatif Adalah Eksponen (Ditandakan*)

Paras Keertian	$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$	
Saiz Sampel Statistik Ujian	10–60	70–100	10–40	50–100	10–40	50–100
D				*		*
MD				*		*
V		*		*		*
W^2		*		*		*
U^2		*		*		*
A^2	*	*	*	*	*	*

RUJUKAN

Naylor, T.S. et al (1966). *Computer Simulation Techniques*. New York: John Wiley.

Pearson K. (1900). *On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably be supposed to have from random sampling*. Philosophical Magazine (5), 50, 157 – 168 1980.

Stephens, M.A. (1974). EdF statistics for Goodness of Fit and Some Comparisions. *JASA*, 69, 730 – 737.

Stephens, M.A. (1970). Use of the Kolmogorov-Smirnov, Cramer-von-Mises and related statistics without extensive tables. *JRSS, Series B*, 32, 115 – 122.